DOCUMENT RESUME

ED 184 619 JC 800 169

AUTHOR Gosser, Jon; And Others
TITLE The Effect of Repeated Testing on Long Term Retention and Generalization in a General Psychology Course.
PUB DATE 26 Jul 79
NOTE 10p.; Paper presented at the Annual Meeting of the American Educational Research Association North Central Region Special Interest Group for Community-Junior College Research (Ann Arbor, MI, July 26, 1979)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Community Colleges; Computer Assisted Testing; Course Objectives; *Educational Testing; *Generalization; Item Banks; *Psychology; Reinforcement; *Retention (Psychology); Scores; Student Attitudes; Teaching Methods; Test Construction; Tests; Two Year Colleges
IDENTIFIERS *Repeated Testing

ABSTRACT

A study involving four sections of a general psychology course at Delta College (Michigan) was conducted to determine: (1) the effectiveness of repeated, cumulative testing on long-term retention of information; (2) its effect on students' ability to generalize information learned as measured by the College Level Examination Program (CLEP) general psychology test; and (3) students' attitudes toward the repeated testing procedure. Classroom tests that were administered during the course of the semester were constructed by computer from a pool of 7,000 test questions that were keyed to specific course objectives. The three sections (57 students) comprising the experimental group were tested in a cumulative fashion, and the control section (19 students) was given non-cumulative tests. To measure long-term retention at the end of the semester, both groups were given the CLEP test and a comprehensive final examination based upon the question pool. Questionnaires were used to assess students' attitudes. Study results showed that while the experimental groups performed significantly better on the final examination composed of questions from the 7,000 item population, there was no significant difference in the CLEP scores of the two groups. Survey responses indicated students' preference for the cumulative testing procedure. (JP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED184619

The Effect of Repeated Testing on Long Term Retention and Generalization in a General Psychology Course

Jon Gosser, Gene Packwood,
and
William Walters

Delta College

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Jon Gosser

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Paper presented at the Annual Meeting of the American Educational Research Association North Central Region Special Interest Group for Community-Junior College Research

Ann Arbor, Michigan
July 26, 1979

JC 800 169

The Effect of Repeated Testing on Long Term Retention and Generalization in a General Psychology Course

July 26, 1979

JON W. GOSSER
Delta College

GENE PACKWOOD
Delta College

WILLIAM M. WALTERS
Delta College

Long term retention was assessed in a general psychology class by comparing the performance of a group of students who were repeatedly tested to that of others who were not. Generalization was tested by comparing scores on the CLEP psychology test. Students attitudes were also assessed. The results were that the repeated testing group performed better on a comprehensive final that was selected from the same 7,000 item population, but the CLEP test scores were not significantly different. Attitude data indicated students' preference for the repeated testing procedure. The results were discussed in light of educational implications.

The Effect of Repeated Testing on Long Term Retention and Generalization in a General Psychology Course

Objectives: The primary objective of the study was to determine if repeated cumulative testing and access to test items from a 7,000 item population would increase long term retention of information in a general psychology course. Long term retention was defined as scores on a comprehensive final examination at the end of the semester. Total test scores and responses to questions that related to the earlier sections of the course (the most frequently retested material) were assessed.

A second objective was to assess the degree of generalization of the information learned by the students as measured by the nationally standardized College Level Examination Program (CLEP) general psychology test.

The third objective was to measure the students' attitudes toward the testing procedure.

Prospective: The theoretical perspective upon which this research is based is the well established effect of positive feedback on learning (see, for example, Logan, 1972). The general phenomenon is that postive feedback aids learning and that more feedback is better than less (Skinner, 1968). Retesting students over the same course materials a number of times and allowing students to keep their tests would seem to be one way of providing this feedback.

A paramound problem, however, exists for the educator who tries to provide adequate feedback for students through retesting. If students are retested using the same questions or from a limited set

of questions, they will be likely, through rote memorization, to learn only a small sample of information. That is, the students will learn specific answers to specific questions. In this way they may do very well on tests made up of these questions or similar questions, but they will not have acquired any general understanding of the concepts being taught.

The standard way of dealing with this narrow focus is to avoid giving students access to specific questions, but only to general non-specific guides concerning the concepts to be tested. Then each test is developed from a large population of items related to those concepts. This approach is utilized by some college professors and many nationally standardized tests such as the Scholastic Aptitude Test, Medical College Admission Test, and CLEP. This approach does eliminate the rote memorization problem and does an adequate job of a summative evaluation of student learning. However, these testing procedures are given at the end of a learning period, and are typically not used as feedback during the learning process itself (formative evaluation).

An alternative approach that combines the benefits of the aforementioned procedure with a feedback system could be used for both formation and summative evaluation of students. The alternative investigated here involves providing students with general objectives or guidelines. Each test is then made up of items from a large population of test questions keyed to the objectives. Then students can be repeatedly tested over the same objectives by using different questions on each test. If this item population is large enough, students will not be able to benefit from rote memorization. Thus students can be provided sample questions for practice, and they can

also keep their tests to study because, given a large enough item pool, they will be unlikely to see the exact question again.

A further issue is the question of generalizability outside of the specific pool of test items. Ideally, a well constructed pool of items will encourage students to continuously review the course materials, and therefore be able to answer additional relevant questions correctly even if they were not part of the original item pool.

Method: A pool of 7,000 test questions was colleged from a variety of sources, primarily test manuals that accompany certain introductory psychology books. Each test item was then keyed to the course objectives for a general psychology class. The specific classroom tests were constructed by computer, using the objective keys.

Four different sections of the course participated in the study. Three sections (57 students) became the experimental group and were tested in a cumulative fashion--each test was comprehensive of the material covered to that point. The control section (19 students) was given non-cumulative tests--they were only tested over material most recently covered. Both classes were given two comprehensive finals at the end of the semester. The first comprehensive final was constructed independently of the instructor from the 7,000 item pool. The second comprehensive was the CLEP test for general psychology. The students' grades in the class would be based on the mean of five classroom tests or either of the comprehensive finals--which ever was higher. In other words, the students had three chances of getting a satisfactory grade in the class--the five classroom tests, the first comprehensive final, and the second comprehensive final (CLEP).

Below is a schematic of the experimental design.

	Retested on All Obj.	Keep Copies of In Class Tests	Given Sample Test-Answers	Sample Tests-Answers in Library	Two Comprehensives*
Experimental	Yes	Yes	Yes	Yes	Yes
Control	No	No	No	No	Yes

*CLEP Exam and Comprehensive Course Exam

As a test of the comparability of the experimental and control groups prior to treatment, scores on the first test and the new material on the second test resulted in no significant differences between the groups.

Data Sources: For long term retention of the test pool material the first comprehensive test (constructed from the 7,000 item test file) scores were compared between experimental and control groups.

Generalization was measured by comparing CLEP test scores (the second comprehensive) between the experimental and control groups.

Students' attitudes were measured by student questionnaire responses for both groups.

Results: T-tests were performed comparing the performance of the experimental group to that of the control. The students' learning was measured by the scores on the comprehensive final examination. Two sets of scores were collected--the portion of the examine which covered information from the earlier portions of the course (the most extensively retested material), and the total comprehensive score.

Table 1

T-tests of Comprehensive Scores

	N	$\bar{X}$	t
Retested Material			
Experimental	57	23.58	
Control	19	20.58	2.33*
Total Comprehensive			
Experimental	57	102.32	
Control	19	92.68	2.005*

*Sig. beyond .05 level

As can been seen from Table 1, the experimental group performed significantly better than the control group on both measures of learning.

As the measure of generalization experimental and controls were compared on CLEP test performance.

Table 2

T-tests of CLEP Scores

	N	$\bar{X}$	t
Experimental	57	46.61	.053 (N.S.)
Control	19	46.74	

As can be seen from Table 2, there were no statistically signficant differences between the experimental and control groups on the measure of generalization.

The attitude scores showed that both groups preferred the retesting procedure. Of the retesting group, 76% found the procedure "is very helpful" or "is helpful" and 64% of the non-retested group rated a description of the procedure in this way.

Summary and Conclusions: Repeated testing increased long term

retention when measured by a comprehensive test taken from the instructor's 7,000 item pool. This is, when students were repeatedly tested over the same concepts, they performed better on a comprehensive test when the comprehensive test items came from the same pool as the other tests.

On the other hand, this improvement did not generalize to a different population of items, as the CLEP test scores for the two groups were not signficantly different. This may be partially due to the fact that the CLEP test scores were not critical for the students' grades, or there was a low relationship between the 7,000 item pool and the CLEP.

The results of this study do not support that repeated testing alone leads to generalization of the students' learning. However, students seemed to prefer the repeated testing procedure.

Educational Importance: Repeated testing seems to be a means of improving long term retention of a finite set of concepts. When one has specific goals and objectives, it may help to repeatedly test students over the same material. This could be useful in courses that ask for recall of factual types of information.

Perhaps methods other than retesting alone would be needed to produce generalization to other, related types of course material not specifically tested during the class. Currently the following are being added to the repeated testing procedure to attempt to increase generalization: (1) provide the page number(s) of the text book where the information needed to answer a given objective can be found to each question, (2) supplementary material on each objective which can either be printed with the test results for questions

missed (the best but an expensive procedure) or made available in the library in order of the class objectives, and (3) improved test questions and related information keyed to class objectives by developing a professionally screen nation-wide bank of questions and related information organized by objective under the direction of the American Pyschological Association's Division on Teaching of Psychology. Whether these additional feedback procedures are able to correct the generalization problem remains to be tested.

References: Logan, Frank A., Learning and Motivation. Dubuque, Iowa: Wm. C. Brown Company Publishers, 1970.

Skinner, B. F., Technology of Teaching. New York: Appleton-Century-Crofts, 1968.

UNIVERSITY OF CALIFORNIA
CLEARINGHOUSE FOR
JUNIOR COLLEGES
96 POWELL LIBRARY BUILDING
LOS ANGELES, CALIFORNIA 90024
EE 36

APR 25 1980